Multiscale Tensor Approximation for Volume Data

Suter, S K ; Zollikofer, C P E ; Pajarola, R

Abstract: Advanced 3D microstructural analysis in natural sciences and engineering depends ever more on modern data acquisition and imaging technologies such as micro-computed or synchrotron tomography and interactive visualization. The acquired high-resolution volume data sets have sizes in the order of tens to hundreds of GBs, and typically exhibit spatially complex internal structures. Such large structural volume data sets represent a grand challenge to be explored, analyzed and interpreted by means of interactive visualization, since the amount of data to be rendered is typically far beyond the current performance limits of interactive graphics systems. As a new approach to tackle this bottleneck problem, we employ higher-order tensor approximations (TAs). We demonstrate the power of TA to represent, and focus on, structural features in volume data. We show that TA yields a high data reduction at competitive rate distortion and that, at the same time, it provides a natural means for multiscale volume feature representation.

Multiscale Tensor Approximation for Volume Data

S. K. Suter[1,2] and C. P. E. Zollikofer[2] and R. Pajarola[1]

[1]Visualization and MultiMedia Lab, University of Zurich, Switzerland
[2]Anthropological Institute und Museum, University of Zurich, Switzerland

Abstract

Advanced 3D microstructural analysis in natural sciences and engineering depends ever more on modern data acquisition and imaging technologies such as micro-computed or synchrotron tomography and interactive visualization. The acquired high-resolution volume data sets have sizes in the order of tens to hundreds of GBs, and typically exhibit spatially complex internal structures. Such large structural volume data sets represent a grand challenge to be explored, analyzed and interpreted by means of interactive visualization, since the amount of data to be rendered is typically far beyond the current performance limits of interactive graphics systems. As a new approach to tackle this bottleneck problem, we employ higher-order tensor approximations (TAs). We demonstrate the power of TA to represent, and focus on, structural features in volume data. We show that TA yields a high data reduction at competitive rate distortion and that, at the same time, it provides a natural means for multiscale volume feature representation.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Viewing algorithms I.4.7 [Computer Graphics]: Feature Measurement—Feature representation

1. Introduction

Non-invasive analysis of organismic structures, tissue and materials with microtomographic techniques has seen a rapid development over the past few years. Micro-computed X-ray tomography (μCT) as now become a standard tool e.g. in clinically-oriented medical research, especially in bone biology. As a relatively recent technology, Synchrotron Tomography (ST) has opened up new areas of research at the sub-micrometer level. Multidimensional microstructural data sets from such advanced 3D scanning technologies regularly reach sizes in the order of tens to hundreds of GB. These large 3D data volumes not only represent an immense amount of information, but also exhibit an increasing level of detail of internal structure in space and time, resulting in a high degree of complexity at different scales. Such data sets typically contain a variety of structures at different levels of scale. These structures are often repetitive, but do not have fractal properties. One of the most recent 3D scanning technologies – phase contrast Synchrotron Tomography (pcST) – has become of special interest for the non-invasive analysis of growth microstructures in hard tissues of living and fossil species [TS08]. For example, pcST permits identification of patterns of daily enamel deposition in fossil hominid teeth, and counting the resulting incremental growth lines on virtual dental cross sections permits an estimate of the age at death of a fossil specimen with an precision of several days. However, current graphics systems do not permit to extract, characterize and visualize the actual growth structures, which typically exhibit a complex three-dimensional pattern [JSM03, MJS03]. There is therefore a need for computer graphics methods which permit full characterization and interactive visualization of internal structures of very large data volumes.

The large amount of 3D data acquired requires extensive preprocessing prior to rendering for visual exploration since the sheer amount of data and the constraints for interactive visualization go well beyond established rendering technologies and exceed commonly available graphics hardware capacities. Hence it is crucial to develop visualization methods that readily adaptable not only to the size but also the structural information contained in the volume data, so as to support interactive frame rates at high display quality. The grand challenge today is thus to make the *implicit* information contained in structural volume data *explicitly* available,

i.e. to render it in an interactive visual form that is readily understandable by the user. Meeting this challenge not only requires further developments in interactive visualization of large data volumes, but it also implies research into new methods of multiscale feature-preserving data reduction.

1.1. Contributions

The contributions of our work are manifold. We demonstrate the power of a tensor approximation (TA) based feature-preserving data reduction method for multidimensional volume data sets. The presented TA approach extracts and reveals the salient volume features at multiple scales (1) and supports progressive and selective reconstruction (2), while at the same time still achieving good rate distortion (3) even compared to state-of-the-art approaches such as wavelet transforms (WT). In contrast to other very successful multiresolution oriented techniques such as WT, which essentially provide progressive accuracy on a corresponding spatial resolution, the TA inherently captures the significant spatial features at different scales but supports their reconstruction at the original or at any sub-sampled spatial resolution. Thus the TA is more appropriate to preserve and highlight structural features independent from the reconstruction resolution, as demonstrated by our experimental results. We provide a proof-of-concept for TA as a suitable mathematical framework for multiscale feature representation in volume data.

2. Related Work

During the last years, there has been an enormous increase in processor speed, memory size and disk storage capacity. However, due to similar technological progress in data acquisition systems, there always exists a gap between the size acquired data and the amount of data that can efficiently be processed in memory. Accordingly, in the context of interactive visualization and exploration there is an ongoing need to reduce and optimize the amount of data to be processed. A fundamental concept of data reduction is to remove redundant and irrelevant information while preserving the most relevant features. Techniques of data reduction are thus directly linked to concepts and techniques of data compression, noise reduction and feature recognition/extraction.

An early data reduction approach has been introduced with the *Fourier Transform*.With the *Wavelet Transform* (WT) any input signal can be decomposed into a localized frequency domain. Today, the WT is a standard method for compressed volume data representation (developed by [Mur93] and refined by e.g., [GEA96, IP98, Rod99, KS99, NS01, GWGS02, SW03]) because it can represent the signal's frequency components more compactly than FT and can be localized in space.

However, the bases retrieved from the classical multiresolution decomposition approaches like WT are obtained regardless of the structures within the data sets. Even tough the bases of WT are adaptively obtained, the WT bases result from convolutions along the three major spatial axes *xyz* with predefined 1D filters, i.e., the compression ratio as well as the bases are limited to these axes. We therefore call these bases *axis-aligned bases*. However, significant features in structural volume data are typically not aligned along axes *xyz*, such that techniques are required which produce *feature-aligned* bases.

In recent years, *tensor-based approximations* (TAs) have become a significant research topic in data compression and visualization [VT04, WWS*05, TS06, WA08, WXC*08, YWT*09]. With TA models, the dimension and size of a data set can be reduced and at the same time the data may be transformed into a compact multilinear representation. In its essence, TA represents an extension of the concept of principal components analysis (PCA) to multi-modal input data. While the PCA computes the data correlations in two modes the TA does it for N modes. With a three-way [KDL80, Kro3c, TBDLK87] or N-way [dLdMV00a, dLdMV00b] analysis, the input data is decomposed into multilinear components consisting of one core tensor and a basis matrix for each mode, the so called Tucker model [Tuc66]. These components store the relevant data, i.e., the direction of the features, and achieve at the same time very high compression ratios. Consequently, the original data is approximated with its most relevant components. The volume can ergo be approximated and represented in its new intrinsic domain representation, which responds to the features contained in the volume data set. The bases in the TA approach are found by taking into account a combination of the matrix correlations along each mode. Such matrix correlations show, in particular when applied with multiple resolutions, the directions of features we look for. That is why we call such bases *feature-aligned bases*.

Furthermore, the WT is typically applied recursively yielding a *multiresolution analysis* decomposition of the input data. As such the WT defines a multiresolution hierarchy of coefficients representing some sort of localized average and detail information. Each coefficient is responsible to improve the approximation of the original data in some well defined and fixed sized spatial region. Hence the WT is particularly designed to minimize the mean scalar differences between approximated and original data at different spatial resolutions, with the higher-level coefficients generally being more important but corresponding to larger locally averaged and blurred data. In contrast, the TA provides an approximation of the original data not at multiple spatial resolutions but instead at multiple scales of feature expressiveness. In addition to previous use of TA in graphics, which focused mainly on image ensembles, we demonstrate the TA's strong feature-preserving characteristic in 3D volumes and analyze its difference to WT, also computationally.

Moreover, we demonstrate that volume reconstruction can be done not only progressively in approximation quality but as well variable in resolution and selective in space. This aspect is important to design interactive visualization systems on TA based volume representations.

Features have been expressed differently in the past. While with PCA the features are expressed in a vector (1^{st}-order tensor), in image ensembles [WA08, YWT*09] the features are represented in a matrix (2^{nd}-order tensor), in our TA-approach, we aim to express the features in three modes (3^{rd}-order tensor), which is, with the tensor decomposition, consisting of three basis matrices and a core tensor. The expressiveness of features in three modes is for some applications more important than mere scalar approximation accuracy.

3. Tensor Decomposed Volume Representation

Our aim is to define a compact feature-preserving data representation in the context of structural volume visualization. The TA framework can achieve this goal through the decomposition of the volume data into a reduced set of bases. Hence the task is to find bases such that the decomposition performs an effective data reduction while retaining all relevant features.

3.1. Tensor Decomposition and Approximation

A tensor is a higher-order generalization of an array: A vector is a 1^{st}-order tensor, a matrix a 2^{nd}-order tensor and a volume is a 3^{rd}-order tensor. In tensor approximation approaches the input data is represented as a higher-order tensor and is decomposed into multilinear components, the so called Tucker model [Tuc66]. This decomposition is usually implemented as an alternating least squares (ALS) algorithm, e.g. *TUCKALS* [KDL80, Kro3c, TBDLK87], a higher-order singular value decomposition (HOSVD) [dLdMV00a], or a higher-order orthogonal iteration (HOOI) [dLdMV00b].

With TA models, the dimension and size of a data set can be reduced and at the same time this data set is transformed into a compact multilinear representation. This works as follows: A N^{th}-order input tensor $\mathcal{A}$ is decomposed into different components with reduced ranks, i.e., one core tensor $\mathcal{B}$ and N basis matrices $\mathbf{U}^{(1)}, \mathbf{U}^{(2)} \dots \mathbf{U}^{(N)}$, so called Tucker tensors [Tuc66]. Then a tensor rank reduction is performed on the Tucker model in a similar way like a matrix rank reduction. The rank reduction is carried out to remove redundancy and to get a coarser level of abstraction of the original data set. Consequently, with TA only the relevant components of the original data set are stored in the tensor decomposition. This tensor decomposition can then be used to reconstruct an approximation $\widetilde{\mathcal{A}}$ of the original data set $\mathcal{A}$. As an additional benefit, a high compression ratio can be achieved by storing the tensor decomposition instead of the original data set.

For volume rendering, the TA can be exploited by selectively reconstructing a reduced-rank volume $\widetilde{\mathcal{A}}$ from its decomposition as required. This can include a specific initial rank-reduced reconstruction level and subsequent progressively higher-rank approximations. Furthermore, as we outlined in Section 4, volume sub-regions can be reconstructed, also at varying spatial resolution.

3.2. Reduced Rank Tensor Decomposition

In the TA approach, a multi-modal data array or tensor $\mathcal{A}$, is approximated by reduced-rank tensors, i.e., the ranks of some core and basis tensors are reduced. For a real N^{th}-order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ of size $\prod_{i=1}^{N} I_i$, a reduced-rank approximation can be defined in two ways:

i A rank-1 approximation is defined as $\widetilde{\mathcal{A}} = \lambda \cdot \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \cdots \otimes \mathbf{U}^{(N)}$ from the tensor-product or the outer-product $\otimes$ of its basis vectors $\mathbf{U}^{(k)} \in \mathbb{R}^{I_k}$ and scalar λ. Hence a tensor $\mathcal{A}$ can also be approximated by a linear combination of rank-1 approximations as $\widetilde{\mathcal{A}} = \sum_{r=1}^{R} \lambda_r \cdot \mathbf{U}_r^{(1)} \otimes \mathbf{U}_r^{(2)} \cdots \otimes \mathbf{U}_r^{(N)}$, which is called a rank-R approximation. This approximation, however, is just a special case of the following definition, which we will adopt in this work. The rank-1 approximation is also known as the parallel factor analysis (PARAFAC) or the canonical decomposition (CANDECOMP).

ii Alternatively, a rank-$(R_1, R_2, \dots R_N)$ approximation of $\mathcal{A}$ is formulated as finding a lower-rank tensor $\widetilde{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ with $rank_n(\widetilde{\mathcal{A}}) = R_n \leq rank_n(\mathcal{A})$, where $rank_n(\mathcal{A})$ is the order of the vector space of the n-mode vectors of $\mathcal{A}$. The n-mode vectors of a multi-modal array, tensor, correspond to the row-vectors or column-vectors of a matrix. We will continue with this more general definition of a reduced-rank tensor approximation. The rank-$(R_1, R_2, \dots R_N)$ approximation is also referred to as the Tucker model.

In general, a reduced-rank approximation is sought such that the least-squares difference is minimized: $\widetilde{\mathcal{A}} = \arg\min(\widetilde{\mathcal{A}}) \parallel \mathcal{A} - \widetilde{\mathcal{A}} \parallel^2$. In the latter approach (ii), the approximated tensor can be represented as $\widetilde{\mathcal{A}} = \mathcal{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)}$ from the n-mode product $\times_n$ of basis tensors in a given reduced rank space. Given that $(R_1, \dots R_N)$ are sufficiently smaller than the initial data dimensions $(I_1, \dots I_N)$, the core tensor $\mathcal{B} \in \mathbb{R}^{R_1 \times R_2 \cdots \times R_N}$ and the basis matrices $\mathbf{U}^{(k)} \in \mathbb{R}^{I_k \times R_k}$ together give rise to a compact approximation $\widetilde{\mathcal{A}}$ of the original tensor $\mathcal{A}$. For a rank-R approximation (i) above, all R_i are equal and the core tensor essentially becomes diagonal.

In Figure 1 we illustrate the (ii) tensor decomposition for a 3^{rd}-order tensor, or volume data set. Note that the three basis matrices $\mathbf{U}^{(k=1\dots3)}$ represent a set of R_k column vectors each

of length I_k, and the core tensor $\mathcal{B}$ is a volume of reduced dimensionality $R_1 \times R_2 \times R_3$.

We can implement $\widetilde{\mathcal{A}} = \mathcal{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$ as a weighted sum over outer vector products

$$\widetilde{\mathcal{A}} = \sum_{r,s,t} \mathcal{B}[r,s,t] \cdot \mathbf{U}^{(1)}_{(r)} \otimes \mathbf{U}^{(2)}_{(s)} \otimes \mathbf{U}^{(3)}_{(t)}, \qquad (1)$$

with $\mathbf{U}_{(r)}$ indicating the r-th column vector of basis matrix $\mathbf{U}$ and the indices r,s,t going from 1 to R_1,R_2,R_3 respectively. Hence for every index combination r,s,t a volume of original size is reconstructed by the outer-products, which is weighted for summation by the corresponding entry in the core tensor.

For more detailed information on reduced-rank tensor approximations and the definitions of n-mode vectors, n-mode products $\times_n$ or outer-products $\otimes$ in the context of tensor, matrices and vectors we refer the reader to [dLdMV00a, dLdMV00b, DLCM08, Kro08, KB09].

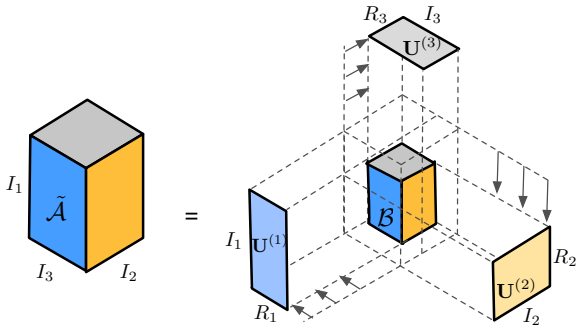


Figure 1: *Visualization of the equation* $\widetilde{\mathcal{A}} = \mathcal{B} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$.

3.3. Multiscale Tensor Approximation

As mentioned before, the goal is to find a reduced-rank tensor $\widetilde{\mathcal{A}}$ that approximates a given input data set $\mathcal{A}$ with as little error $\varepsilon = \sum_{r,s,t} (\mathcal{A}[r,s,t] - \widetilde{\mathcal{A}}[r,s,t])^2$ as possible. The root mean squared error (RMSE) $\frac{1}{I_1 \cdot I_2 \cdot I_3} \sqrt{\varepsilon}$ can then be analyzed in a rate-distortion sense dependent on the number of (non-zero) coefficients of the reduced-rank tensor decomposition, as will be shown in Section 5.

To compute the tensor approximations, we decompose the volume into a Tucker decomposition. As we demonstrate in Section 5, different rank-$(R_1,R_2,\ldots R_N)$ approximations yield reconstructions of the original data at different scales of feature expressiveness and RMSE. Therefore, we call a Tucker decomposition at varying ranks $(R_1,R_2,\ldots R_N)$ a multiscale approach. In particular, we show in Section 4.3 that different lower rank approximations can be defined progressively from one initial higher rank-$(R_1,R_2,\ldots R_N)$ approximation. Thus indeed one single tensor decomposition leads to a multiscale TA representation.

4. Reconstruction

The basic process to reconstruct a volume $\widetilde{\mathcal{A}}$ from its decomposition $\mathcal{B}$ and $\mathbf{U}^{(1..3)}$ is given in Equation 1. Given an original volume of size $I_1 \cdot I_2 \cdot I_3$ this involves summation of $R_1 \cdot R_2 \cdot R_3$ such volumes. However, to improve reconstruction efficiency in practice, the outer-products can be computed incrementally by rearranging the summation as follows

$$\widetilde{\mathcal{A}} = \sum_r \mathbf{U}^{(1)}_{(r)} \otimes \sum_s \mathbf{U}^{(2)}_{(s)} \otimes \sum_t \mathcal{B}[r,s,t] \cdot \mathbf{U}^{(3)}_{(t)}, \qquad (2)$$

resulting in a significantly reduced complexity of only $O(R_1 \cdot I_1 \cdot I_2 \cdot I_3)$. The order of the modes can be permuted such that R_1 has in fact the smallest rank.

The total space cost of the reduced rank representation consists of $R_1 \cdot R_2 \cdot R_3 + R_1 \cdot I_1 + R_2 \cdot I_2 + R_3 \cdot I_3$ coefficients only, in contrast to the $I_1 \cdot I_2 \cdot I_3$ size of the original volume.

4.1. Spatial Selection

The reconstruction from a reduced rank tensor decomposition can be localized and restricted in space along any of the original data dimensions $I_{1..3}$. As indicated in Figure 2, a selective subspace of the 3D volume can be reconstructed by limiting the computation in Equations 1 and 2 to the corresponding subranges of the basis matrices $\mathbf{U}^{(k)}$ in the index dimensions I_k respectively.

Therefore, for any given subspace $J_1 \times J_2 \times J_3 \subseteq I_1 \times I_2 \times I_3$ we can construct the reduced-rank basis matrices $\widehat{\mathbf{U}}^{(k)} \in \mathbb{R}^{J_k \times R_k}$ by only selecting the $J_k \subseteq I_k$ rows from the original bases $\mathbf{U}^{(k)}$. With these rank-reduced bases $\widehat{\mathbf{U}}^{(k)}$ we can then perform a partial spatial reconstruction $\widehat{\mathcal{A}}$. See also Section 5 for examples.

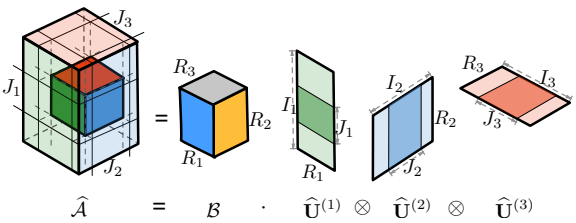


Figure 2: *Reconstruction of a spatially localized tensor approximation* $\widehat{\mathcal{A}}$ *from rank-reduced basis matrices* $\widehat{\mathbf{U}}^{(k)}$.

4.2. Variable Resolution

As shown above, for spatial selection the reconstruction can be reduced in the data dimensions I_k. However, not only can the data dimensions I_k be rangebound for spatial selectivity, but they can also be subsampled and interpolated for varying spatial resolution as outlined in Figure 3. Hence instead of selecting a compact subrange of rows we can derive a set of

reduced-resolution basis matrices $\overline{\mathbf{U}}^{(k)} \in \mathbb{R}^{J_k \times R_k}$ by subsampling the original bases $\mathbf{U}^{(k)}$ at any reduced rate $J_k < I_k$ in the corresponding index dimensions. A lower resolution volume $\overline{\mathcal{A}}$ can then be reconstructed from these reduced-resolution bases.

Various re-sampling and interpolation methods can be considered to define the reduced-resolution bases $\overline{\mathbf{U}}^{(k)}$. However, simple regular subsampling achieves very good results as demonstrated in Section 5. Furthermore, similar to mipmapping [Wil83], reduced resolution matrix pyramids could be generated to form a multiresolution hierarchy of TAs.

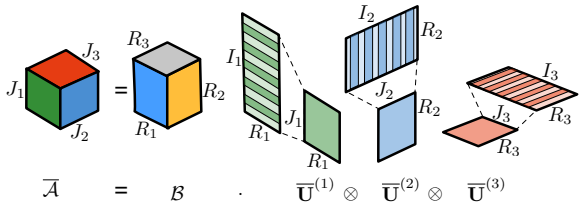


Figure 3: *A reduced resolution tensor approximation $\overline{\mathcal{A}}$ from reduced-resolution basis matrices $\overline{\mathbf{U}}^{(k)}$.*

4.3. Progressive Reconstruction

The data dimensions I_k define the original spatial extent and resolution of the data set and can be used for variable spatial reconstruction as outlined above. After the HOSVD, the tensor decomposition according to Equation 1 exhibits in its rank space a natural order, where the ranks with a higher impact on the data set are generally ordered first. Consequently, an initial high-quality tensor approximation with large rank dimensions $(R_1, \ldots R_N)$ yields a truly integrated multiscale tensor approximation that allows variable rank reconstruction using any index subranges $r_k \in 1 \ldots \leq R_k$.

Nevertheless, the given rank order from the HOSVD decomposition is not the only exclusive solution to find all properties/features in a data set. In fact, the order of the column vectors of each basis $\mathbf{U}^{(k)}$ could be shifted, corresponding to exchanging k-mode planes of coefficients in the core tensor $\mathcal{B}$ accordingly. For example, the basis column vectors could be ordered according to their accumulated factors in $\mathcal{B}$: for $\mathbf{U}^{(1)}_{(r)}$ we compute its weight as $\sum_{s,t} \mathcal{B}[r,s,t]$ and then permutate the column vectors of $\mathbf{U}^{(1)}$ in decreasing order.

Note that since the data dimensions I_k and the rank space R_k are independent of each other, spatial selection or resampling in I_k can be combined orthogonally with progressive reconstruction in R_k. Hence a spatially selective multiresolution and multiscale reconstruction can be defined using the proposed rank-reduced tensor decomposition, approximation and reconstruction framework.

5. Experimental Results

5.1. Experimental Setup

For the application of TA to structural volume data, there are three phases: (a) definition of the reduced ranks (R_1, R_2, R_3), (b) decomposition of the volume $\mathcal{A}$ into a core tensor $\mathcal{B}$ and three basis matrices $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}$, and $\mathbf{U}^{(3)}$, and (c) reconstruction of the volume $\widetilde{\mathcal{A}}$ from the reduced rank tensor decomposition. For our experiments in this paper, an HOOI generating a rank-(R_1, R_2, R_3) approximation was used. In order to achieve various reduction levels, a series of different rank-(R_1, R_2, R_3) approximations are compared.

The choice of ranks chosen for different tensor approximations is often done experimentally (comparing the approximations visually). As a rule of thumb for our experiments we reduce the ranks for each compression/approximation level by an additional factor of two for every dimension I_N. I.e., we have chosen the ranks according to the following scheme: For the first compression ratio, the ranks (R_1, R_2, R_3) were chosen at half the size of the original dimensions, i.e., for a volume with the dimensions $I_1 \times I_2 \times I_3$ the ranks were set to $R_k = \frac{I_k}{2}$; and for each iteration of rank reduction the R_k were further divided by 2. The same function of ranks has already been applied by others authors, e.g. [WXC*08].

Note that in the present experiments we perform the tensor decomposition as well as the volume reconstruction in Matlab using the *Matlab Tensor Classes* [BK06, BK07, BK, KS08], which achieve the Tucker decomposition by implementing the HOSVD [dLdMV00a] and HOOI [dLdMV00b]. The initial guess for the Tucker ALS is computed on eigenvalues. Therefore, while we can demonstrate the reduction and approximation power of TA and its applicability to feature-preserving multiscale volume representation, the implementation and experiments in Matlab do not already lead to a system applicable to multi-GB volume data sets. Hence moderately sized models have been used to demonstrate the capability of the outlined TA approach.

5.2. Applications

We assess the performance of the new TA approach with two different types of volume data sets. First, we use the bonsai tree example to assess how TA performs compared to state-of-the-art methods (see Section 5.3), especially regarding feature detection versus compression efficiency. Second, we test the performance of TA in rendering periodic features at various spatial scales, using as examples synchrotron microstructural data of tooth enamel and synthetic data simulating periodic microstructures (Section 5.4). The bonsai tree represents an object with conspicuous features at multiple scales: trunk, branches, and leaves. An effective multiscale approach should be able to characterize these different objects as a function of the data reduction level. Ideally, the

specific features should turn on and off on their adequate respective levels.

We have compared the data reduction and the associated feature extraction ability of tensor approximation to the state-of-the-art, i.e., wavelet transform. Specifically, we show multilevel Haar wavelet reconstructions. While other wavelets may generated smoother reconstructions (Figure 4), such wavelets nevertheless suffer from the same general problem of being least-squares optimized instead of feature sensitive. Thus the original data is reconstructed as best as possible at different resolutions but not specifically at different scales of features.

(a) WT (b) WT (c) WT (d) WT (e) TA

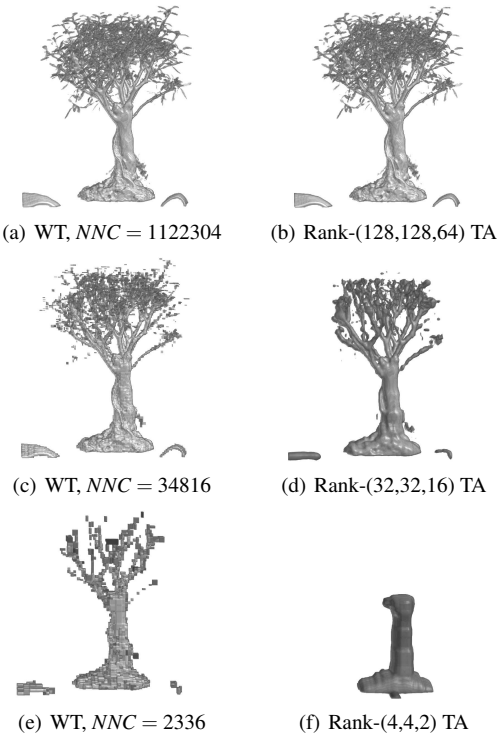
Figure 4: *Comparison of wavelets at NNC* $= 311296$*: (a) Haar WT, (b) Biorthogonal 9/3 WT, (c) Daubechies3 WT, (d) discrete Meyer WT, (e) rank-(64,64,64) TA.*

5.3. Feature Expressiveness

Figure 5 shows the comparison of TA and a seven-level WT at corresponding compression ratios. For each reconstruction, the same number of non-zero coefficients (*NNC*) has been chosen. In TA, the $NNC = R_1 \cdot R_2 \cdot R_3 + R_1 \cdot I_1 + R_2 \cdot I_2 + R_3 \cdot I_3$ represents the number of coefficients of the tensor decomposition, while in WT, we select the *NNC* most significant coefficients. A high-quality approximation at half of the original ranks for TA and corresponding number of WT coefficients, shows comparable reconstructions in Figure 5(a) and 5(b). At lower rank reconstructions, however, the TA is able to extract the larger scale features such as tree trunk and branches specifically, while the WT still describes all features but at coarser resolutions. While the WT might achieve an optimal approximation in a rate distortion sense, the TA shows a more specific multiscale feature expressiveness dependent on the reduction level. An interpretation of these results together with the analysis of the rate distortion is given in Section 6.

5.4. Volume Microstructures

The dental enamel data set is interesting here because it represents periodic growth structures that occur at different levels of scale, and exhibit different spatial orientations. Human tooth enamel has a microstructure that is roughly comparable to a bunch of densely packed spaghetti (see simulated dental structures after [JSM03, MJS03] in Figure 7(b)). During dental enamel formation, each "spaghetto" (so-called dental enamel prism) elongates in centrifugal direction through the daily apposition of a small segment of



(a) WT, *NNC* $= 1122304$ (b) Rank-(128,128,64) TA

(c) WT, *NNC* $= 34816$ (d) Rank-(32,32,16) TA

(e) WT, *NNC* $= 2336$ (f) Rank-(4,4,2) TA

Figure 5: *The* $256^2 \times 128$ *bonsai tree reconstructed from different compression levels. (a,c,e) Show WT reconstructions. (b,d,f) Show TA of different rank-*(R_1, R_2, R_3) *reductions. TA achieves much better feature selectivity at different scales (trunk, branches, leafs) dependent on data reduction levels than WT.*

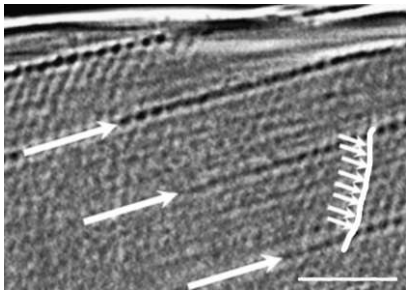
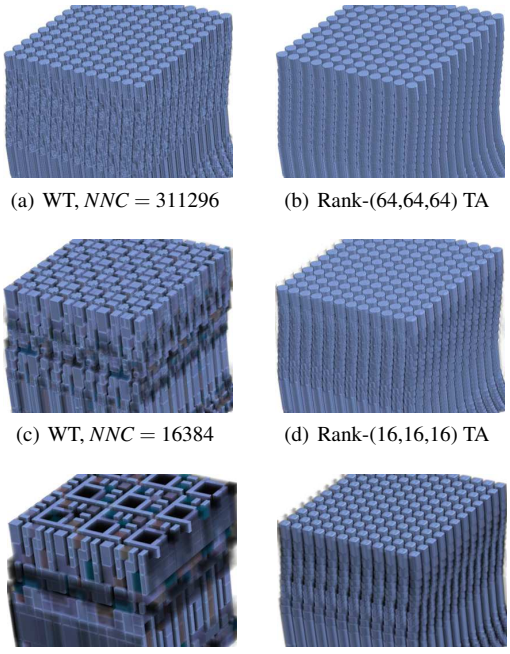


Figure 6: *Cross-sectional image of tooth enamel (pcST data; scale = 50 microns). Small arrows: cross striations; large arrows: Retzius lines. The direction of the growth prisms is orthogonal to the cross striations. From [TS08].*

enamel. Daily growth increments are visible as surfaces perpendicular to the longitudinal direction of the prisms (so-called cross striations). In addition, approximately weekly growth halts are visible as so-called Retzius lines. As shown in Figure 6, the spatial scale and orientation of these structures is highly characteristic for each feature. This volume data set thus represents an ideal test case to assess the per-

formance of TA regarding the use of axis-aligned versus feature-aligned basis decompositions. As outlined in Section 1.1, the latter method is proposed here to perform two tasks simultaneously: to compress very large volume data efficiently and to extract the biologically relevant multiscale features. We first test the performance of TA with a synthetic data set simulating dental growth structures as a set of tubes slightly twisted in two dimensions. As shown in Figure 7 even an extremely compact rank-(4,4,4) TA preserves the periodic features, while in the WT the features disintegrate and merge at the corresponding reduction level. Moreover, the WT approximation in Figure 7(e) shows that the microstructures can only be reconstructed with a strong bias of axis alignment.



(a) WT, $NNC = 311296$ (b) Rank-(64,64,64) TA

(c) WT, $NNC = 16384$ (d) Rank-(16,16,16) TA

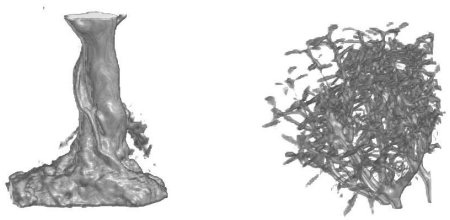
(e) WT, $NNC = 3136$ (f) Rank-(4, 4, 4) TA

Figure 7: *Simulated periodic microstructures (256^3, 8bit voxel depth): (a,c,e) Different approximation levels from eight-level Haar WT, and (b,d,f) different rank-(R_1,R_2,R_3) TA with corresponding NNC. In contrast to WT, TA achieves a feature-preserving reconstruction with very few coefficients (3136).*

Results of TA application to real microstructural data of tooth enamel are shown in Figure 8 and 9. Figure 8 demonstrates that a compact TA (rank-(2,2,2); 392 coefficients) permits to highlight features (growth prisms) that are difficult to identify and visualize in the original data set (Figure 8(a)–8(d)). The corresponding *NNC* of the WT reconstruction cannot characterize any features (Figure 8(e)–8(h)). Figure 9 demonstrates TA-based multiscale feature extraction. Different rank-(R_1,R_2,R_3) approximation levels highlight the specific periodic microstructures explained in Figure 6. TA thus has great potential to be a highly effective shape-preserving feature detector at multiple scales.

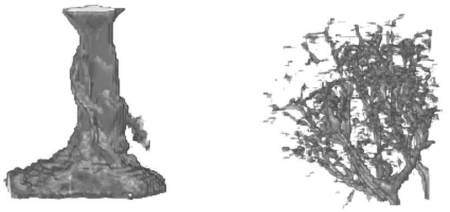
5.5. Selective Reconstruction

With the visualizations in Figures 10 and 11 we demonstrate the selective reconstruction of a subsample of row vectors of the basis matrices as described in the Sections 4.1 and 4.2. In Figure 10(a) we show the selective reconstruction for the trunk with a subset of the basis matrices vectors, using $\mathbf{U}^{(1)}_{(80\ldots190)}, \mathbf{U}^{(2)}_{(150\ldots256)}$ and $\mathbf{U}^{(3)}_{(15\ldots90)}$. In Figure 10(b) the subset of basis matrices vectors to selectively reconstruct some branches have been chosen as $\mathbf{U}^{(1)}_{(1\ldots128)}, \mathbf{U}^{(2)}_{(1\ldots128)}$ and $\mathbf{U}^{(3)}_{(1\ldots128)}$. The same concept, only using a sub-set of the basis matrices vectors for a lower resolution reconstruction, has been applied to perform a subsampling of the volume (Figure 11).



(a) (b)

Figure 10: *Selective parts of the $256^2 \times 128$ bonsai tree reconstructed from a rank-$(128,128,64)$ TA and with different subsets of basis matrices row vectors. Either (a) the trunk or (b) the branches could be selectively reconstructed.*

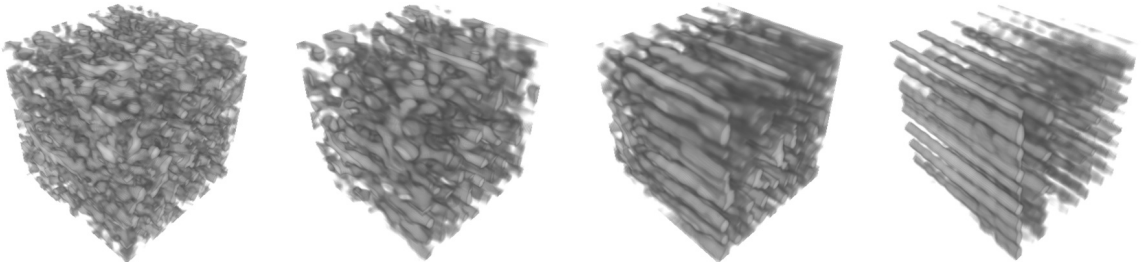


(a) (b)

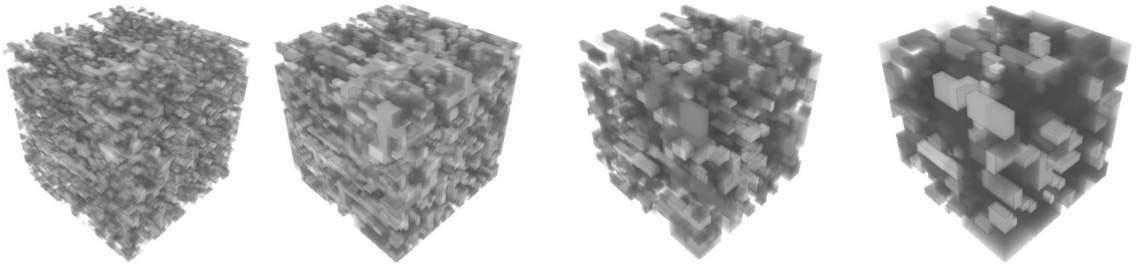
Figure 11: *Subsampling: (a) for the volume in Figure 10(a) and (b) for the volume in Figure 10(a).*

5.6. Rate Distortion

The performance of TA versus WT can also be compared in terms of rate distortion (Figure 12), i.e., measuring the root mean-square error (RMSE) of each approximation relative to the original data set. For the simulated periodic growth structures, the rate distortion for TA was significantly better than for WT. However, in all the other examples, the rate distortion curves for corresponding WT and TA are either

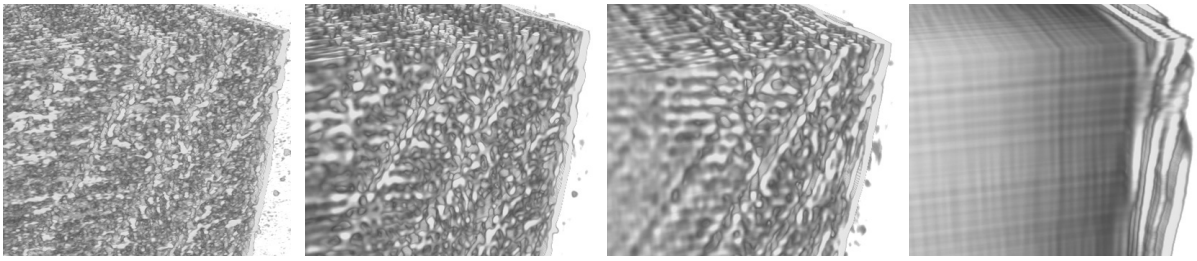


(a) Rank-(32,32,32) TA (b) Rank-(16,16,16) TA (c) Rank-(8,8,8) TA (d) Rank-(2,2,2) TA



(e) WT, $NNC = 38912$ (f) WT, $NNC = 7168$ (g) WT, $NNC = 2048$ (h) WT, $NNC = 392$

Figure 8: *Structural volume data of tooth enamel acquired with phase contrast synchrotron tomography (64^3 voxels, 16bit voxel depth, 0.75 microns resolution per voxel). (a-d) reconstructions from four different rank reductions of a tensor decomposition; (e-h) eight-level Haar wavelet reconstructed volumes at corresponding compression levels (= the same number of non-zero coefficients). Note that TA achieves high compression ratios, and simultaneously extracts/highlights relevant features (growth prisms) in the data set by moving from a high-ranked to a low-ranked tensor decomposition. WT does not recover these features even at moderate compression rates.*



(a) Original (b) Rank-(64,64,64) TA (c) Rank-(32,32,32) TA (d) Rank-(8,8,8) TA

Figure 9: *Periodic microstructures in tooth enamel (256^3, 16bit voxel depth, 0.75 microns per voxel). (a) original data set; (b)-(d) feature extraction/visualization with different rank-(R_1,R_2,R_3) approximations (on the front side of the volume, growth prisms are oriented left-to-right; Retzius lines are oriented bottom-left to top-right).*

close to each other or WT is slightly better. These results can be interpreted as follows: RMSE is an optimal measure of goodness-of-fit for data sets under the hypothesis that features exhibit a random distribution. The WT method performs well regarding this criterion, but it cannot recover features exhibiting non-random distributions (see Figures 7–9). The TA method, on the other hand, performs well in terms of non-random feature extraction; this results in slightly worse RMSE values, because random information is filtered out.

5.7. Performance

The complexity for the inverse discrete WT in three modes is $O(C \cdot I_1 \cdot I_2 \cdot I_3)$, where the constant C depends on the filter F and the number of levels L of a multilevel WT, $C = L \cdot 6 \cdot F \cdot (F-1)$. Note that this computational complexity estimate does not yet include the computational costs for the e.g. run length encoding to eliminate the thresholded wavelet coefficients. The complexity for TA would require $O(R_1 \cdot R_2 \cdot R_3 \cdot I_1 \cdot I_2 \cdot I_3)$ operations in a straightforward implementation. However, the reconstruction complexity can be optimized and hence reduced to only $O(R_1 \cdot I_1 \cdot I_2 \cdot I_3)$

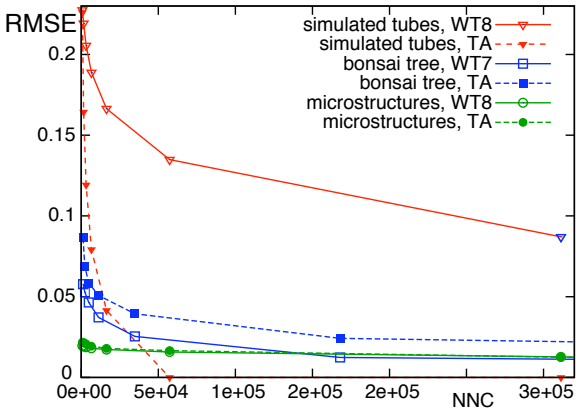


Figure 12: *Rate distortion for TA and WT<levels> reconstructions (see Figures 5, 7, 9).*

(see Section 4). Both, the complexity for WT and TA grow with increasing volume size. The complexity for WT also increases with the squared filter size, while the complexity for TA depends on the smallest achievable rank. In-depth investigations on the complexity of TA and its potential regarding a sufficient rank reduction has been made in [KK07].

6. Interpretation of Experiments

As has been shown earlier in the graphics field [WWS*05, WXC*08], at large compression ratios, TA can generate higher quality images than WT or principal components analysis (PCA). Here, we went one step further and showed that the mathematical framework of TA not only permits to generate high-quality images at high compression rates, but also permits to quantify, characterize, and visualize non-random features in structural volume data. Such features can be highlighted by using different approximation/ compression levels for the reconstruction. Classical approaches such as the WT do not achieve comparable results. We also showed that classical measures of compression quality, such as rate distortion diagrams, are of limited use to assess the quality of feature extraction methods. Here, new procedures need to be implemented, which permit comparative quantitative analysis of the performance of compression-based feature detectors.

As exposed in the Introduction, the TA approach proposed here represents a first step toward the implementation of new tools for interactive rendering of very large volume data sets in the range of 100s GB. The path toward such a goal can be sketched as follows: (A) Here, we demonstrated the feature-preserving capabilities of TA at high compression rates. (B) Extension to large volumes can proceed via volume bricking techniques. Bricking is required because the reconstruction of the tensor decomposition increases in cubic order. At the same time, bricking permits out-of-core techniques to load very large data sets (similar approaches have already been proposed in [WWS*05, WXC*08]), and it is an ideal precondition for parallelizing the TA computational workload. Further research into parallel, hierarchical and out-of-core TA methods is thus necessary to achieve extension to the 100-GB-volume rendering domain.

7. Conclusion

New data acquisition techniques lead to volume data sets of ever-increasing size, which tend to be one step ahead of the available graphics sources for interactive visualization. There is thus an ongoing need to develop new data compression/feature extraction methods to tackle the resulting performance bottlenecks. This paper represents a proof-of-concept study demonstrating that TA methods represent a powerful progressive approach to (a) represent microstructural volume data sets at high compression ratios, and (b) simultaneously extract relevant features at different spatial scales. We have shown that, compared to classical approaches such as WT, the TA approach is especially appropriate to concomitantly perform compression/feature extraction, because the tensor decompositions have feature-aligned bases, i.e., they represent at least three-dimensional spatial correlations of the structures of interest. We conclude that TA has a great potential as a progressive multi-scale feature-preserving volume approximations approach. Finally, we provide a scenario of how the specific capabilities of TA can be used to tackle the "very large volume rendering problem".

Acknowledgments

The authors wish to thank volvis.org for the bonsai tree volume model and [TS08] for the image in Figure 6. We also want to thank Philipp Schlegel for some help on the visualization of example data sets. This work was supported in part by a grant from Forschungskredit 2009 of the University of Zurich, Switzerland.

References

[BK] BADER B. W., KOLDA T. G.: Matlab tensor toolbox version 2.3, http://csmr.ca.sandia.gov/ tgkolda/tensortoolbox/. 5

[BK06] BADER B. W., KOLDA T. G.: Algorithm 862: Matlab tensor classes for fast algorithm prototyping. *ACM Trans. Math. Softw. 32*, 4 (2006), 635–653. 5

[BK07] BADER B. W., KOLDA T. G.: Efficient matlab computations with sparse and factored tensors. *Siam Journal On Scientific Computing 30*, 1 (2007), 205–231. 5

[DLCM08] DE LATHAUWER L., COMON P., MASTRONARDI N.: Special issue on tensor decompositions and applications. *SIAM Journal On Matrix Analysis and Applications 30*, 3 (2008), VII–VII. 4

[dLdMV00a] DE LATHAUWER L., DE MOOR B., VANDEWALLE J.: A multilinear singular value decomposition. *SIAM Journal of Matrix Analysis and Applications 21*, 4 (2000), 1253–1278. 2, 3, 4, 5

[dLdMV00b] DE LATHAUWER L., DE MOOR B., VANDEWALLE J.: On the best rank-1 and rank-($R_1, R_2, ..., R_N$) approximation of higher-order tensors. *SIAM Journal of Matrix Analysis and Applications 21*, 4 (2000), 1324–1342. 2, 3, 4, 5

[GEA96] GROSSO R., ERTL T., ASCHOFF J.: Efficient data structures for volume rendering of wavelet-compressed data. In *Winter School of Computer Graphics* (1996), Computer Society Press. 2

[GWGS02] GUTHE S., WAND M., GONSER J., STRASSER W.: Interactive rendering of large volume data sets. In *Proceedings IEEE Visualization* (2002), Computer Society Press, pp. 53–60. 2

[IP98] IHM I., PARK S.: Wavelet-based 3D compression scheme for very large volume data. In *Graphics Interface* (June 1998), pp. 107–116. 2

[JSM03] JIANG Y., SPEARS I. R., MACHO G. A.: An investigation into fractured surfaces of enamel of modern human teeth: a combined sem and computer visualisation study. *Arch Oral Biol 48*, 6 (Jun 2003), 449–457. 1, 6

[KB09] KOLDA T. G., BADER B. W.: Tensor decompositions and applications. *Siam Review 51*, 3 (Sept. 2009), 455–500. 4

[KDL80] KROONENBERG P., DE LEEUW J.: Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika 45* (1980), 69–97. 2, 3

[KK07] KHOROMSKIJ B. N., KHOROMSKAIA V.: Low rank tucker-type tensor approximation to classical potentials. *Central European Journal of Mathematics 5*, 3 (SEP 2007), 523–550. 9

[Kro08] KROONENBERG P. M.: *Applied Multiway Data Analysis*. Wiley's Series in Probability and Statistics, January 2008. 4

[Kro3c] KROONENBERG P. M.: *Three-mode principal component analysis: Theory and applications*. Leiden: DSWO Press, 1983c. 2, 3

[KS99] KIM T.-Y., SHIN Y.-G.: An efficient wavelet-based compression method for volume rendering. In *Proceedings Pacific Graphics 99* (1999), pp. 147–156. 2

[KS08] KOLDA T. G., SUN J.: Scalable tensor decompositions for multi-aspect data mining. In *ICDM* (2008), IEEE Computer Society, pp. 363–372. 5

[MJS03] MACHO G. A., JIANG Y., SPEARS I. R.: Enamel microstructure–a truly three-dimensional structure. *J Hum Evol 45*, 1 (Jul 2003), 81–90. 1, 6

[Mur93] MURAKI S.: Volume data and wavelet transform. *IEEE Computer Graphics and Applications 13*, 4 (July 1993), 50–56. 2

[NS01] NGUYEN K. G., SAUPE D.: Rapid high quality compression of volume data for visualization. In *Proceedings EUROGRAPHICS* (2001), pp. 49–56. also in Computer Graphics Forum 20(3). 2

[Rod99] RODLER F.: Wavelet based 3D compression with fast random access for very large volume data. In *Proceedings Pacific Graphics* (1999), pp. 108–117. 2

[SW03] SCHNEIDER J., WESTERMANN R.: Compression domain volume rendering. In *Proceedings IEEE Visualization* (2003), pp. 293–300. 2

[TBDLK87] TEN BERGE J. M. F., DE LEEUW J., KROONENBERG P. M.: Some additional results on principal components analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika 52* (1987), 183–191. 2, 3

[TS06] TSAI Y.-T., SHIH Z.-C.: All-frequency precomputed radiance transfer using spherical radial basis functions and clustered tensor approximation. *ACM Transactions on Graphics 25*, 3 (2006), 967–976. 2

[TS08] TAFFOREAU P., SMITH T. M.: Nondestructive imaging of hominoid dental microstructure using phase contrast x-ray synchrotron microtomography. *Journal of Human Evolution Article in press* (2008). 1, 6, 9

[Tuc66] TUCKER L. R.: Some mathematical notes on three-mode factor analysis. *Psychometrika 31*, 3 (September 1966), 279–311. 2, 3

[VT04] VASILESCU M. A. O., TERZOPOULOS D.: Tensortextures: multilinear image-based rendering. *ACM Trans. Graph. 23*, 3 (2004), 336–342. 2

[WA08] WANG H., AHUJA N.: A tensor approximation approach to dimensionality reduction. *International Journal of Computer Vision 76*, 3 (2008), 217–229. 2, 3

[Wil83] WILLIAMS L.: Pyramidal parametrics. In *Proceedings ACM SIGGRAPH* (1983), ACM SIGGRAPH, pp. 1–11. 5

[WWS*05] WANG H., WU Q., SHI L., YU Y., AHUJA N.: Out-of-core tensor approximation of multi-dimensional matrices of visual data. *ACM Transactions on Graphics 24*, 3 (August 2005), 527–535. 2, 9

[WXC*08] WU Q., XIA T., CHEN C., LIN H.-Y. S., WANG H., YU Y.: Hierarchical tensor approximation of multidimensional visual data. *IEEE Transactions on Visualization and Computer Graphics 14*, 1 (January/February 2008), 186–199. 2, 5, 9

[YWT*09] YAN S., WANG H., TU J., TANG X., HUANG T. S.: Mode-kn factor analysis for image ensembles. *IEEE Trans Image Process 18*, 3 (Mar 2009), 670–676. 2, 3